

1 PROCESS FOR DIAGNOSIS OF PHYSIOLOGICAL CONDITIONS BY

2 CHARACTERIZATION OF PROTEOMIC MATERIALS

3
4 FIELD OF THE INVENTION

5 This invention generally relates to the use of proteomic
6 investigation as a diagnostic tool; and particularly to the
7 use of proteomic investigative techniques and methodology to
8 determine a proteomic basis for the development and
9 progression of abnormal physiological conditions.

10
11 BACKGROUND OF THE INVENTION

12 At the present time there exist numerous diagnostic
13 techniques and procedures whose goal is to assess an
14 individual's physiological condition. From a very early age,
15 individuals are subjected to a variety of routine physical
16 examinations with the goal of maintaining a vibrant and
17 healthful existence. During the course of these
18 examinations, a physician will often require a variety of
19 diagnostic procedures based upon several factors, for example
20 the patient's physical presentation, familial history,
21 environmental factors which may place the patient at
22 particular risk, and tests to ascertain or predict the course
23 or progress of known conditions.

24 Routine tests generally include blood and urine analysis

and X-rays, and often include electrocardiogram (EKG), cardiac stress tests and the like. Dependent upon preliminary findings, additional tests may be ordered, in accordance with current standards of care, and may include computer assisted tomography (CAT) scans, magnetic resonance imagery (MRI), echocardiographic studies, Doppler analysis, angiograms, elctromyograph (EMG), electroencephelograph (EEG), and the like procedures which are geared to assist the physician in forming a definitive diagnosis. The majority of these tests are directed toward quantifying a particular condition, usually during a point of exacerbation of the condition.

Unfortunately, even the most skilled diagnostician may not always be able to successfully determine the reasons for a particular clinical condition or the underlying cause of the manifestation of certain symptoms. Thus, conditions are often misdiagnosed, and medications are often ordered which are inappropriate or ineffective. Furthermore, very few tests exist which offer the diagnostician a prospective method of analyzing the propensity for an individual to develop a particular condition.

As we delve more deeply into our genetic makeup, we are becoming increasingly aware of genetic anomalies which cause us to be particularly inclined to either develop or manifest

1 a wide variety of conditions. The genetic information of all
2 living organisms (e.g. animals, plants and microorganisms) is
3 encoded in deoxyribonucleic acid (DNA). In humans, the
4 complete genome is now believed to be comprised of about
5 30,000 - 40,000 genes located on 24 chromosomes.

6 While each of these genes, or nucleotide sequences,
7 encodes a single protein, or several splice variants
8 (approximately 10 or more) these may be post-translationally
9 modified into many different forms having different molecular
10 masses. Subsequent to their expression via transcription,
11 translation, and post-translational modification, each
12 protein or fragment thereof is capable of fulfilling a
13 specific biochemical function within a living cell.

14 Changes in a DNA sequence are known as mutations and can
15 result in proteins with altered or in some cases even lost
16 biochemical activities; this in turn can cause genetic
17 disease. Such mutations may include nucleotide deletions,
18 insertions or alterations (i.e. point mutations). Point
19 mutations can be either "missense", resulting in a change in
20 the amino acid sequence of a protein or "nonsense" coding for
21 a stop codon and thereby leading to a truncated protein.

22 It is currently believed that there are more than 3000
23 genetically related diseases including hemophilias,
24 thalassemias, Duchenne Muscular Dystrophy (DMD), Huntington's

Disease (HD), Alzheimer's Disease and Cystic Fibrosis (CF). In addition to mutated genes, which result in genetic disease, certain birth defects are the result of chromosomal abnormalities such as Trisomy 21 (Down's Syndrome), Trisomy 13 (Patau Syndrome), Trisomy 18 (Edward's Syndrome), Monosomy X (Turner's Syndrome) and other sex chromosome aneuploidies such as Klinefelter's Syndrome (XXY). Further, there is growing evidence that certain DNA sequences may predispose an individual to any of a number of diseases such as diabetes, arteriosclerosis, obesity, various autoimmune diseases and cancer (e.g. colorectal, breast, ovarian, lung).

The science of proteomics recognizes that messenger RNAs, which are transcripts of genomic DNA that directly encode proteins, are assemblable in a variety of ways, and that expressed proteins can further be modified, e.g. by methods such as phosphorylation and glycosylation leading to variations in protein expression.

As broadly defined, leading experts in the field of proteomics describe the science as including transcriptional profiling to determine those genes which are transcribed into RNA in a particular cell type, developmental stage or disease state. The science seeks to provide methods and techniques for high-throughput expression and purification of proteins.

Additionally, the science of proteomics seeks to study

1 protein profiling by the use of various techniques, so-called
2 proteomic investigative techniques, including two-dimensional
3 gel electrophoresis and mass spectroscopy,
4 co-immunoprecipitation, affinity chromatography, protein
5 binding analysis, overlay analysis and BIACORE, use of the
6 yeast two-hybrid method for studies of protein-protein
7 interaction, pathway analysis for interpreting signal
8 transduction and complex cellular processes, three-
9 dimensional structure studies and large-scale protein
10 folding, and the use of bioinformatics analysis of proteomics
11 data.

12 Various techniques have been put forth for analyzing the
13 protein constituents of either whole cells or of cell
14 organelles. By separating proteins in a first dimension
15 based upon charge and in a second dimension based upon
16 molecular size, individual proteins on the gel can be
17 isolated and characterized. The drawback of this technique is
18 that the gels are difficult to analyze, their resolving power
19 often being insufficient to separate the various distinct
20 proteins present in a particular sample. Furthermore, there
21 is a distinct lack of reproducibility from one gel sample to
22 the next and the paucity of reproducible data along with the
23 similarity of data between different tissues, species and
24 organism states makes the development of bioinformatics

1 databases problematic. Although there have been advances in
2 software techniques to bring greater degrees of
3 standardization and reproducibility to 2D-gel analysis,
4 significant obstacles remain.

5 There is ongoing research in the field of protein
6 expression profiling using 2D-gel in conjunction with other
7 techniques. Using laser capture microdissection, researchers
8 obtain both diseased and normal cells. Using 2D-gel
9 analysis, all the protein components in these cells are
10 separated and capillary high performance liquid
11 chromatography (HPLC) or electrospray ion-trap mass
12 spectroscopy are utilized to identify differing levels of
13 protein expression in diseased versus normal cells.

14 An additional technique in proteomics is the use of
15 phage display, wherein peptide or protein libraries are
16 created on viral surfaces and are then screened on a mass
17 scale. Since the proteins remain with their encoding genes,
18 identification is facilitated. This is more valuable as a
19 genomics tool than a proteomics tool since differential
20 expression is still not usefully elucidated. A similar
21 technique called profusion forms molecules which are
22 conjugates in which a peptide or protein is chemically linked
23 to its encoding mRNA, therefore facilitating affinity
24 screening techniques. In addition, techniques exist for

1 identifying antibody fragments which bind human proteins.
2 Detection is simplified by tagging each antibody fragment
3 with a peptide encoding sequence. Subsequent testing of
4 tissue samples for the presence of corresponding target
5 proteins can then be studied so as to determine their
6 relevance as possible therapeutic or diagnostic agents.

7 Methods utilizing mass spectrometry for the analysis of
8 a target polypeptide have been taught wherein the polypeptide
9 is first solubilized in an appropriate solution or reagent
10 system. The type of solution or reagent system, e.g.,
11 comprising an organic or inorganic solvent, will depend on
12 the properties of the polypeptide and the type of mass
13 spectrometry performed and are well known in the art (see,
14 e.g., Vorm et al. (1994) Anal. Chem. 66:3281 (for MALDI) and
15 Valaskovic et al. (1995) Anal. Chem. 67:3802 (for ESI). Mass
16 spectrometry of peptides is further disclosed, e.g., in WO
17 93/24834 by Chait et al.

18 In one prior art embodiment, the solvent is chosen so
19 that the risk that the molecules may be decomposed by the
20 energy introduced for the vaporization process is
21 considerably reduced, or even fully excluded. This can be
22 achieved by embedding the sample in a matrix, which can be an
23 organic compound, e.g., sugar, in particular pentose or
24 hexose, but also polysaccharides such as cellulose. These

1 compounds are decomposed thermolytically into CO₂ and H₂O so
2 that no residues are formed which might lead to chemical
3 reactions. The matrix can also be an inorganic compound,
4 e.g., nitrate of ammonium which is decomposed practically
5 without leaving any residues. Use of these and other solvents
6 are further disclosed in U.S. Pat. No. 5,062,935 by Schlag et
7 al.

8 Prior art mass spectrometer formats for use in analyzing
9 the translation products include ionization (I) techniques,
10 including but not limited to matrix assisted laser desorption
11 (MALDI), continuous or pulsed electrospray (ESI) and related
12 methods (e.g., IONSPRAY or THERMOSPRAY), or massive cluster
13 impact (MCI); these ion sources can be matched with detection
14 formats including linear or non-linear reflection time-of-
15 flight (TOF), single or multiple quadropole, single or
16 multiple magnetic sector, Fourier Transform ion cyclotron
17 resonance (FTICR), ion trap, and combinations thereof (e.g.,
18 ion-trap/time-of-flight). For ionization, numerous
19 matrix/wavelength combinations (MALDI) or solvent
20 combinations (ESI) can be employed. Subattomole levels of
21 protein have been detected, for example, using ESI
22 (Valaskovic, G. A. et al., (1996) Science 273:1199-1202) or
23 MALDI (Li, L. et al., (1996) J. Am. Chem. Soc. 118:1662-1663)
24 mass spectrometry.

1 ES mass spectrometry has been introduced by Fenn et al.
2 (J. Phys. Chem. 88, 4451-59 (1984); PCT Application No. WO
3 90/14148) and current applications are summarized in recent
4 review articles (R. D. Smith et al., Anal. Chem. 62, 882-89
5 (1990) and B. Ardrey, Electrospray Mass Spectrometry,
6 *Spectroscopy Europe*, 4, 10-18 (1992)). MALDI-TOF mass
7 spectrometry has been introduced by Hillenkamp et al.
8 ("Matrix Assisted UV-Laser Desorption/Ionization: A New
9 Approach to Mass Spectrometry of Large Biomolecules,"
10 *Biological Mass Spectrometry* (Burlingame and McCloskey,
11 editors), Elsevier Science Publishers, Amsterdam, pp. 49-60,
12 1990). With ESI, the determination of molecular weights in
13 femtomole amounts of sample is very accurate due to the
14 presence of multiple ion peaks which all could be used for
15 the mass calculation.

16 The mass of the target polypeptide determined by mass
17 spectrometry is then compared to the mass of a reference
18 polypeptide of known identity. In one embodiment, the target
19 polypeptide is a polypeptide containing a number of repeated
20 amino acids directly correlated to the number of
21 trinucleotide repeats transcribed/translated from DNA; from
22 its mass alone the number of repeated trinucleotide repeats
23 in the original DNA which coded it, may be deduced.

24 U.S. Patent No. 6,020,208 utilizes a general category of

1 probe elements (i.e., sample presenting means) with Surfaces
2 Enhanced for Laser Desorption/Ionization (SELDI), within
3 which there are three (3) separate subcategories. The SELDI
4 process is directed toward a sample presenting means (i.e.,
5 probe element surface) with surface-associated (or surface-
6 bound) molecules to promote the attachment (tethering or
7 anchoring) and subsequent detachment of tethered analyte
8 molecules in a light-dependent manner, wherein the said
9 surface molecule(s) are selected from the group consisting of
10 photoactive (photolabile) molecules that participate in the
11 binding (docking, tethering, or crosslinking) of the analyte
12 molecules to the sample presenting means (by covalent
13 attachment mechanisms or otherwise).

14 PCT/EP/04396 teaches a process for determining the
15 status of an organism by peptide measurement. The reference
16 teaches the measurement of peptides in a sample of the
17 organism which contains both high and low molecular weight
18 peptides and acts as an indicator of the organism's status.
19 The reference concentrates on the measurement of low
20 molecular weight peptides, i.e. below 30,000 Daltons, whose
21 distribution serves as a representative cross-section of
22 defined controls. Contrary to the methodology of the instant
23 invention, the '396 patent strives to determine the status of
24 a healthy organism, i.e. a "normal" and then use this as a

1 reference to differentiate disease states. The present
2 inventors do not attempt to develop a reference "normal", but
3 rather strive to specify particular markers which are
4 evidentiary of at least one specific disease state, whereby
5 the presence of said marker serves as a positive indicator of
6 disease. This leads to a simple method of analysis which can
7 easily be performed by an untrained individual, since there
8 is a positive correlation of data. On the contrary, the '396
9 patent requires a complicated analysis by a highly trained
10 individual to determine disease state versus the perception
11 of non-disease or normal physiology.

12 Richter et al, Journal of Chromatography B, 726(1999)
13 25-35, refer to a database established from human
14 hemofiltrate comprised of a mass database and a sequence
15 database. The goal of Richter et al was to analyze the
16 composition of the peptide fraction in human blood. Using
17 MALDI-TOF, over 20,000 molecular masses were detected
18 representing an estimated 5,000 different peptides. The
19 conclusion of the study was that the hemofiltrate (HF)
20 represented the peptide composition of plasma. No
21 correlation of peptides with relation to normal and/or
22 disease states is made.

23 As used herein, "analyte" refers to any atom and/or
24 molecule; including their complexes and fragment ions. In the

1 case of biological molecules/macromolecules or "biopolymers",
2 such analytes include but are not limited to: proteins,
3 peptides, DNA, RNA, carbohydrates, steroids, and lipids.
4 Note that most important biomolecules under investigation for
5 their involvement in the structure or regulation of life
6 processes are quite large (typically several thousand times
7 larger than H₂O).

8 As used herein, the term "molecular ions" refers to
9 molecules in the charged or ionized state, typically by the
10 addition or loss of one or more protons (H⁺).

11 As used herein, the term "molecular fragmentation" or
12 "fragment ions" refers to breakdown products of analyte
13 molecules caused, for example, during laser-induced
14 desorption (especially in the absence of added matrix).

15 As used herein, the term "solid phase" refers to the
16 condition of being in the solid state, for example, on the
17 probe element surface.

18 As used herein, "gas" or "vapor phase" refers to
19 molecules in the gaseous state (i.e., in vacuo for mass
20 spectrometry).

21 As used herein, the term "analyte desorption/ionization"
22 refers to the transition of analytes from the solid phase to
23 the gas phase as ions. Note that the successful
24 desorption/ionization of large, intact molecular ions by

1 laser desorption is relatively recent (circa 1988)--the big
2 breakthrough was the chance discovery of an appropriate
3 matrix (nicotinic acid).

4 As used herein, the term "gas phase molecular ions"
5 refers to those ions that enter into the gas phase. Note that
6 large molecular mass ions such as proteins (typical
7 mass=60,000 to 70,000 times the mass of a single proton) are
8 typically not volatile (i.e., they do not normally enter into
9 the gas or vapor phase). However, in the procedure of the
10 present invention, large molecular mass ions such as proteins
11 do enter the gas or vapor phase.

12 As used herein in the case of MALDI, the term "matrix"
13 refers to any one of several small, acidic, light absorbing
14 chemicals (e.g., nicotinic or sinapinic acid) that is mixed
15 in solution with the analyte in such a manner so that, upon
16 drying on the probe element, the crystalline matrix-embedded
17 analyte molecules are successfully desorbed (by laser
18 irradiation) and ionized from the solid phase (crystals) into
19 the gaseous or vapor phase and accelerated as intact
20 molecular ions. For the MALDI process to be successful,
21 analyte is mixed with a freshly prepared solution of the
22 chemical matrix (e.g., 10,000:1 matrix:analyte) and placed on
23 the inert probe element surface to air dry just before the
24 mass spectrometric analysis. The large fold molar excess of

1 matrix, present at concentrations near saturation,
2 facilitates crystal formation and entrapment of analyte.

3 As used herein, "energy absorbing molecules (EAM)"
4 refers to any one of several small, light absorbing chemicals
5 that, when presented on the surface of a probe, facilitate
6 the neat desorption of molecules from the solid phase (i.e.,
7 surface) into the gaseous or vapor phase for subsequent
8 acceleration as intact molecular ions. The term EAM is
9 preferred, especially in reference to SELDI. Note that
10 analyte desorption by the SELDI process is defined as a
11 surface-dependent process (i.e., neat analyte is placed on a
12 surface composed of bound EAM). In contrast, MALDI is
13 presently thought to facilitate analyte desorption by a
14 volcanic eruption-type process that "throws" the entire
15 surface into the gas phase. Furthermore, note that some EAM
16 when used as free chemicals to embed analyte molecules as
17 described for the MALDI process will not work (i.e., they do
18 not promote molecular desorption, thus they are not suitable
19 matrix molecules).

20 As used herein, "probe element" or "sample presenting
21 device" refers to an element having the following properties:
22 it is inert (for example, typically stainless steel) and
23 active (probe elements with surfaces enhanced to contain EAM
24 and/or molecular capture devices).

1 As used herein, "MALDI" refers to Matrix-Assisted Laser
2 Desorption/Ionization As used herein, "TOF" stands for Time-
3 of-Flight.

4 As used herein, "MS" refers to Mass Spectrometry.

5 As used herein "MALDI-TOF MS" refers to Matrix-assisted
6 laser desorption/ionization time-of-flight mass spectrometry.

7 As used herein, "ESI" is an abbreviation for
8 Electrospray ionization.

9 As used herein, "chemical bonds" is used simply as an
10 attempt to distinguish a rational, deliberate, and
11 knowledgeable manipulation of known classes of chemical
12 interactions from the poorly defined kind of general
13 adherence observed when one chemical substance (e.g., matrix)
14 is placed on another substance (e.g., an inert probe element
15 surface). Types of defined chemical bonds include
16 electrostatic or ionic (+/-) bonds (e.g., between a
17 positively and negatively charged groups on a protein
18 surface), covalent bonds (very strong or "permanent" bonds
19 resulting from true electron sharing), coordinate covalent
20 bonds (e.g., between electron donor groups in proteins and
21 transition metal ions such as copper or iron), and
22 hydrophobic interactions (such as between two noncharged
23 groups).

24 As used herein, "electron donor groups" refers to the

case of biochemistry, where atoms in biomolecules (e.g, N, S, O) "donate" or share electrons with electron poor groups (e.g., Cu ions and other transition metal ions).

With the advent of mass spectroscopic methods such as MALDI and SELDI, researchers have begun to utilize a tool that holds the promise of uncovering countless biopolymers which result from translation, transcription and post-translational transcription of proteins from the entire genome.

Operating upon the principles of retentate chromatography, SELDI MS involves the adsorption of proteins, based upon their physico-chemical properties at a given pH and salt concentration, followed by selectively desorbing proteins from the surface by varying pH, salt, or organic solvent concentration. After selective desorption, the proteins retained on the SELDI surface, the "chip", can be analyzed using the CIPHERGEN protein detection system, or an equivalent thereof. Retentate chromatography is limited, however, by the fact that if unfractionated body fluids, e.g. blood, blood products, urine, saliva, and the like, along with tissue samples, are applied to the adsorbent surfaces, the biopolymers present in the greatest abundance will compete for all the available binding sites and thereby prevent or preclude less abundant biopolymers from

1 interacting with them, thereby reducing or eliminating the
2 diversity of biopolymers which are readily ascertainable.

3 If a process could be devised for maximizing the
4 diversity of biopolymers discernable from a sample, the
5 ability of researchers to accurately determine the relevance
6 of such biopolymers with relation to one or more disease
7 states would be immeasurably enhanced. Such determinations
8 would then lead to the production of protein expression
9 profiles. These profiles or phenomic fingerprints may be
10 used to simultaneously monitor multiple protein markers
11 associated with differing biological states.

12 What is therefore lacking in the art is a rapid process
13 for separation of proteomics materials, which are variously
14 defined as an "analyte" referring to any atom and/or
15 molecule; including their complexes and fragment ions; or in
16 the case of biological molecules/macromolecules or
17 "biopolymers", wherein such materials include but are not
18 limited to: proteins, peptides, DNA, RNA, carbohydrates,
19 steroids, and lipids, polypeptides, peptide fragments,
20 modified proteins, non-limiting examples of which are
21 glycoproteins, lipoproteins and the like, and related
22 cellular and sub-cellular components. Additionally lacking
23 is a method for identification of their function as it
24 relates to either a normal or an abnormal physiological

1 state, and a method for comparing the presence or absence of
2 particular proteomic materials or groupings thereof, in
3 living cells, which would be indicative or predictive of the
4 presence or predicted development of an abnormal
5 physiological condition or state.

6
7 DESCRIPTION OF THE PRIOR ART

8 United States Patent 5,010,175 discloses a method for
9 producing and selecting peptides with specific properties
10 comprising obtaining selected individual peptides or families
11 thereof which have a target property and optionally
12 determining the amino acid sequence of a selected peptide or
13 peptides to permit synthesis in practical quantities.

14 United States Patent 5,538,897 teaches a method for
15 correlating a peptide fragment mass spectrum with amino acid
16 sequences derived from a database. A peptide is analyzed by a
17 tandem mass spectrometer to yield a peptide fragment mass
18 spectrum. A protein sequence database or a nucleotide
19 sequence database is used to predict one or more fragment
20 spectra for comparison with the experimentally derived
21 fragment spectrum. In one embodiment, sub-sequences of the
22 sequences found on the database which define a peptide having
23 a mass substantially equal to the mass of the peptide
24 analyzed by the tandem mass spectrometer are identified as

1 candidate sequences. For each candidate sequence, a plurality
2 of fragments of the sequence are identified and the masses
3 and m/z ratios of the fragments are predicted and used to
4 form a predicted mass spectrum. The various predicted mass
5 spectra are compared to the experimentally derived fragment
6 spectrum using a closeness-of-fit measure, preferably
7 calculated with a two-step process, including a calculation
8 of a preliminary score and, for the highest-scoring predicted
9 spectra, calculation of a correlation function. While useful
10 to determine the source of a particular fragment, the method
11 fails to teach or suggest a means for diagnosing a
12 physiological condition by characterization of proteomic
13 materials.

14 U.S. Patent 5,808,300 teaches that MALDI MS has been used
15 to generate images of samples in one or more pictures,
16 providing the capability of mapping concentrations of
17 specific molecules in X,Y coordinates of the original sample.
18 For sections of mammalian tissue, for example, this can be
19 accomplished in two ways. First, tissue slices can be
20 directly analyzed after thorough drying and application of a
21 thin coating of matrix by electrospray. Second, imprints of
22 the tissue can be analyzed by blotting the dry tissue
23 sections on specially prepared targets, e.g., C-18 beads.
24 Peptides and small proteins bind to the C-18 and create a

1 positive imprint of the tissue which can be imaged by MALDI
2 MS after application of matrix. Such images can be displayed
3 in individual m/z values as a selected ion image which would
4 localize individual compounds in the tissue, as summed ion
5 images, or as a total ion image which would be analogous to a
6 photomicrograph. This imaging process may also be applied to
7 separation techniques where a physical track or other X,Y
8 deposition process is utilized, for example, in the CE/MALDI
9 MS combination where a track is deposited on a membrane
10 target.

11 U.S. Patent 6,043,031 provides fast and highly accurate
12 mass spectrometer based processes for detecting a particular
13 nucleic acid sequence in a biological sample. Depending on
14 the sequence to be detected, the processes can be used, for
15 example, to diagnose a genetic disease or chromosomal
16 abnormality; a predisposition to a disease or condition,
17 infection by a pathogenic organism, or for determining
18 identity or heredity.

19 U. S. Patent 6,189,013 discloses a project-based full
20 length biomolecular sequence database which is a relational
21 database system for storing biomolecular sequence information
22 in a manner that allows sequences to be catalogued and
23 searched according to association with one or more projects
24 for obtaining full-length biomolecular sequences from shorter

1 sequences. The relational database has sequence records
2 containing information identifying one or more projects to
3 which each of the sequence records belong. Each project
4 groups together one or more biomolecular sequences generated
5 during work to obtain a full-length gene sequence from a
6 shorter sequence. The computer system has a user interface
7 allowing a user to selectively view information regarding one
8 or more projects. The relational database also provides
9 interfaces and methods for accessing and manipulating and
10 analyzing project-based information.

11

12 SUMMARY OF THE INVENTION

13 The instantly disclosed invention is drawn to a process
14 for determining a proteomic basis, e.g. a basis for
15 diagnosing the existence of or predicting the development
16 and/or progression of abnormal physiological conditions based
17 upon the presence of proteomic materials, by first
18 obtaining a patient sample containing such proteomic
19 material(s); preparing said patient sample to facilitate
20 proteomic investigation thereof; isolating one or more
21 patient specific proteomic materials from said patient
22 sample; and comparing said one or more isolated patient
23 specific proteomic materials against a library of proteomic
24 materials having characteristics identifiable with both

1 normal and abnormal physiological conditions or predictive
2 hallmarks thereof. The proteomic materials may be separated
3 into desired sets of diverse moieties by the use of one or
4 more preparations steps. This process permits analysis of
5 one or more of these isolated patient specific proteomic
6 materials thereby enabling the diagnostician to ultimately
7 characterize an individual's condition as being either
8 positively or negatively indicative of one or more abnormal
9 physiological conditions or predictive hallmarks thereof.

10 Also disclosed is a process for sequencing said one or
11 more isolated patient specific proteomic materials, wherein
12 the particular peptide/polypeptide, proteins, nucleotide or
13 oligonucleotide, or the like proteomic material associated
14 therewith is identified. This information permits the
15 development of quantifiable data-linking methodologies upon
16 the appreciation of particular proteomic materials with
17 particular physiological abnormalities.

18 As a useful diagnostic tool, the process of the
19 invention further includes the step of developing at least
20 one antibody to said isolated patient specific proteomic
21 material and may subsequently express at least one protein
22 marker specific to said at least one antibody to said
23 isolated patient specific proteomic material.

24 As a means of determining the significance of an

1 isolated proteomic material, the process may include at least
2 one interactive mapping step to characterize said material.
3 The interactive mapping step may include one or more steps
4 selected from the group consisting of creation of engineered
5 antibodies or proteins, directly determining the three-
6 dimensional structure of said antibody or protein directly
7 from an amino acid sequence thereof; cellular localization,
8 sub-cellular localization, protein-protein interaction,
9 receptor-ligand interaction, and pathway delineation.
10 Included in such mapping techniques may be co-
11 immunoprecipitation, protein or antibody affinity
12 chromatography, protein binding analysis including BIACORE,
13 U.V. spectra, overlay analysis, far Western analysis, immuno-
14 metric analysis, and ELISA. As referred to in this
15 disclosure, engineered antibodies or proteins include, but
16 are not limited to, those which are tagged with a material
17 selected from the group consisting of GFP, colloidal gold,
18 streptavidin, avidin and biotin. Proteomic materials are
19 illustrated by, but not limited to, proteins, peptides or
20 fragments thereof and related isomers and retro-isomers, e.g.
21 an immunologically reactive/detectable fragment thereof,
22 glycoproteins, lipoproteins, modified proteins and the like,
23 antibodies and protein marker.

24 Accordingly, it is an objective of the instant invention

1 to teach methods for proteomic investigation.

2 It is another objective of the instant invention to
3 define a particularly isolated proteomic material which is
4 useful in evidencing and categorizing at least one particular
5 physiological condition or predictive hallmark thereof.

6 It is another objective of the instant invention to
7 evaluate samples containing a plurality of
8 analytes/biopolymers for the presence of physiological
9 condition specific sequences.

10 It is a further objective of the instant invention to
11 elucidate essentially all biopolymeric moieties contained
12 therein, whereby particularly significant moieties may be
13 identified.

14 It is a further objective of the instant invention to
15 provide at least one purified antibody which is specific to
16 said particularly isolated proteomic material.

17 It is yet another objective of the instant invention to
18 teach a monoclonal antibody which is specific to said
19 particularly isolated proteomic material.

20 It is a still further objective of the invention to
21 teach polyclonal antibodies raised against said particularly
22 isolated proteomic material.

23 It is yet an additional objective of the instant
24 invention to teach a diagnostic kit for determining the

1 presence of said particularly isolated proteomic material.

2 It is a still further objective of the instant invention
3 to teach methods for characterizing disease state based upon
4 the identification of said particularly isolated proteomic
5 material.

6 Other objects and advantages of this invention will
7 become apparent from the following description taken in
8 conjunction with the accompanying drawings wherein are set
9 forth, by way of illustration and example, certain
10 embodiments of this invention. The drawings constitute a
11 part of this specification and include exemplary embodiments
12 of the present invention and illustrate various objects and
13 features thereof.

14 BRIEF DESCRIPTION OF THE FIGURES

15 Figure 1 is a block diagram illustrating the proteomic
16 investigative process.

17 DETAILED DESCRIPTION OF THE INVENTION

18 Serum samples from individuals were analyzed using
19 Surface Enhanced Laser Desorption Ionization (SELDI) as a
20 proteomic investigative technique using the Ciphergen
21 PROTEINCHIP system. The chip surfaces included, but were not
22 limited to IMAC-3-Ni, SAX2 surface chemistries, gold chips,
23 and the like.

24 Preparatory to the conduction of the SELDI MS procedure,

various preparatory steps were carried out in order to maximize the diversity of discernible moities educable from the sample. Utilizing a type of micro-chromatographic column called a C18-ZIPTIP available from the Millipore company, the following preparatory steps were conducted.

1. Dilute sera in sample buffer
2. Aspirate and dispense ZIP TIP in 50% Acetonitrile
3. Aspirate and dispense ZIP TIP in Equilibration solution
4. Aspirate and Dispense in serum sample
5. Aspirate and Dispense ZIP TIP in Wash solution
6. Aspirate and Dispense ZIP TIP in Elution Solution

Illustrative of the various buffering compositions useful in the present invention are:

Sample Buffers (various low pH's): Hydrochloric acid (HCl), Formic acid, Trifluoroacetic acid (TFA),

Equilibration Buffers (various low pH's): HCl, Formic acid, TFA;

Wash Buffers (various low pH's): HCl, Formic acid, TFA;

Elution Solutions (various low pH's and % Solvents):

HCl, Formic acid, TFA;

Solvents: Ethanol, Methanol, Acetonitrile.

Spotting was then performed, for example upon a Gold Chip in the following manner:

- 1 1. spot 2 ul of sample onto each spot
2 2. let sample partially dry
3 3. spot 1 ul of matrix, and let air dry.

4 **HiQ Anion Exchange Mini Column Protocol**

- 5 1. Dilute sera in sample/running buffer;
6 2. Add HiQ resin to column and remove any air bubbles;
7 3. Add Ul water to aid in column packing;
8 1. Add sample/running buffer to equilibrate column;
9 2. Add diluted sera;
10 3. Collect all the flow through fraction in Eppendorf
11 tubes until level is at resin;
12 4. Add sample/running buffer to wash column;
13 5. Add elution buffer and collect elution in Eppendorf
14 tubes.

15 Illustrative of the various buffering compositions useful in
16 this technique are:
17 Sample/Running buffers: including but not limited to Bicine
18 buffers of various molarities, pH's, NaCl content, Bis-Tris
19 buffers of various molarities, pH's, NaCl content,
20 Diethanolamine of various molarities, pH's, NaCl content,
21 Diethylamine of various molarities, pH's, NaCl content,
22 Imidazole of various molarities, pH's, NaCl content, Tricine
23 of various molarities, pH's, NaCl content, Triethanolamine of
24 various molarities, pH's, NaCl content, Tris of various

1 molarities, pH's, NaCl content.
 2 Elution Buffer: Acetic acid of various molarities, pH's,
 3 NaCl content, Citric acid of various molarities, pH's, NaCl
 4 content, HEPES of various molarities, pH's, NaCl content, MES
 5 of various molarities, pH's, NaCl content, MOPS of various
 6 molarities, pH's, NaCl content, PIPES of various molarities,
 7 pH's, NaCl content, Lactic acid of various molarities, pH's,
 8 NaCl content, Phosphate of various molarities, pH's, NaCl
 9 content, Tricine of various molarities, pH's, NaCl content.

10 **Chelating Sepharose Mini Column**

- 11 1. Dilute Sera in Sample/Running buffer;
- 12 2. Add Chelating Sepharose slurry to column and allow
- 13 column to pack;
- 14 3. Add UF water to the column to aid in packing;
- 15 4. Add Charging Buffer once water is at the level of the
- 16 resin surface;
- 17 5. Add UF water to wash through non bound metal ions once
- 18 charge buffer washes through;
- 19 6. Add running buffer to equilibrate column for sample
- 20 loading;
- 21 7. Add diluted serum sample;
- 22 8. Add running buffer to wash unbound protein;
- 23 9. Add elution buffer and collect elution fractions for
- 24 analysis;

1 10. Acidify each elution fraction.
2 Illustrative of the various buffering compositions useful in
3 this technique are:
4 Sample/Running buffers including but not limited to Sodium
5 Phosphate buffers at various molarities and pH's;
6 Charging buffers including but not limited to Nickel
7 Chloride, Nickel Sulphate, Copper II Chloride, Zinc Chloride
8 or any suitable metal ion solution;
9 Elution Buffers including but not limited to Sodium
10 phosphate buffers at various molarities and pH's containing
11 various molarities of EDTA and/or Imidazole.

12 **HiS Cation Exchange Mini Column Protocol**

13 1. Dilute sera in sample/running buffer;
14 2. Add HiS resin to column and remove any air bubbles;
15 3. Add Uf water to aid in column packing;
16 4. Add sample/running buffer to equilibrate column for sample
17 loading;
18 5. Add diluted sera to column;
19 6. Collect all flow through fractions in Eppendorf tubes
20 until level is at resin.
21 7. Add sample/running buffer to wash column.
22 8. Add elution buffer and collect elution in Eppendorf
23 tubes.

1 Illustrative of the various buffering compositions useful in
2 this technique are:

3 Sample/Running buffers: including but not limited to Bicine
4 buffers of various molarities, pH's, NaCl content, Bis-Tris
5 buffers of various molarities, pH's, NaCl content,
6 Diethanolamine of various molarities, pH's, NaCl content,
7 Diethylamine of various molarities, pH's, NaCl content,
8 Imidazole of various molarities, pH's, NaCl content, Tricine
9 of various molarities, pH's, NaCl content, Triethanolamine of
10 various molarities, pH's, NaCl content, Tris of various
11 molarities, pH's, NaCl content.

12 Elution Buffer: Acetic acid of various molarities, pH's,
13 NaCl content, Citric acid of various molarities, pH's, NaCl
14 content, HEPES of various molarities, pH's, NaCl content, MES
15 of various molarities, pH's, NaCl content, MOPS of various
16 molarities, pH's, NaCl content, PIPES of various molarities,
17 pH's, NaCl content, Lactic acid of various molarities, pH's,
18 NaCl content, Phosphate of various molarities, pH's, NaCl
19 content, Tricine of various molarities, pH's, NaCl content.

20 The procedure for profiling serum samples is described below:

21 Following the preparatory steps illustrated above,
22 various methods for use of the PROTEINCHIP arrays, available
23 for purchase from Ciphergen Biosystems (Palo Alto, CA), may
24 be practiced. Illustrative of one such method is as follows.

1 The first step involved treatment of each spot with 20 ml of
2 a solution of 0.5 M EDTA for 5 minutes at room temperature in
3 order to remove any contaminating divalent metal ions from
4 the surface. This was followed by rinsing under a stream of
5 ultra-filtered, deionized water to remove the EDTA. The
6 rinsed surfaces were treated with 20 ml of 100 mM Nickel
7 sulfate solution for 5 minutes at room temperature after
8 which the surface was rinsed under a stream of ultra-
9 filtered, deionized water and allowed to air dry.
10 Serum samples (2 ml) were applied to each spot (now "charged"
11 with the metal-Nickel) and the PROTEINCHIP was returned to
12 the plastic container in which it was supplied. A piece of
13 moist KIMWIPE was placed at the bottom of the container to
14 generate a humid atmosphere. The cap on the plastic tube was
15 replaced and the chip allowed to incubate at room temperature
16 for one hour. At the end of the incubation period, the chip
17 was removed from the humid container and washed under a
18 stream of ultra-filtered, deionized water and allowed to air
19 dry. The chip surfaces (spots) were now treated with an
20 energy-absorbing molecule that helps in the ionization of the
21 proteins adhering to the spots for analysis by Mass
22 Spectrometry. The energy-absorbing molecule in this case was
23 sinapinic acid and a saturated solution prepared in 50%
24 acetonitrile and 0.05% TFA was applied (1 ml) to each spot.

1 The solution was allowed to air dry and the chip analyzed
2 immediately using MS (SELDI).

3 Serum samples from patients suffering from a variety of
4 disease states were analyzed using one or more protein chip
5 surfaces, e.g. a gold chip or an IMAC nickel chip surface as
6 described above and the profiles were analyzed to discern
7 notable sequences which were deemed in some way evidentiary
8 of at least physiological condition or disease state.

9 Patient specific samples were obtained and the data used
10 to formulate a library of proteomic materials having
11 characteristics identifiable with both normal and abnormal
12 physiological conditions or predictive hallmarks thereof.
13 Data which is exemplary of the information retrieved via the
14 novel proteomic investigative techniques of the instant
15 invention is set forth in Appendix A.

16 Although all manner of biomarkers related to all
17 disease conditions are deemed to be within the purview of the
18 instant invention and methodology, particular significance
19 was given to those markers and diseases associated with the
20 complement system and Syndrome X and diseases related
21 thereto.

22 The complement system is an important part of
23 non-clonal or innate immunity that collaborates with acquired
24 immunity to destroy invading pathogens and to facilitate the

1 substrate of component C1s and is activated when C1s
2 hydrolyzes a small fragment (C4a) from the amino terminus of
3 the α chain, exposing a binding site on the larger fragment
4 (C4b).

5 The native C3 component consists of two polypeptide
6 chains, α and β . As a serum protein, C3 is involved in the
7 alternative pathway. Serum C3, which contains an unstable
8 thioester bond, is subject to slow spontaneous hydrolysis
9 into C3a and C3b. The C3f component is involved in the
10 regulation required of the complement system which confines
11 the reaction to designated targets. During the regulation
12 process, C3b is cleaved into two parts: C3bi and C3f. C3bi
13 is a membrane-bound intermediate wherein C3f is a free
14 diffusible (soluble) component.

15 Complement components have been implicated in the
16 pathogenesis of several disease conditions. C3 deficiencies
17 have the most severe clinical manifestations, such as
18 recurrent bacterial infections and immune-complex diseases,
19 reflecting the central role of C3. The rapid profusion of
20 C3f moieties and resultant "accidental" lysis of normal cells
21 mediated thereby gives rise to a host of auto-immune
22 reactions. The ability to understand and control these
23 mechanisms, along with their attendant consequences, will

1 enable practitioners to develop both diagnostic and
2 therapeutic avenues by which to thwart these maladies.

3 In the course of defining a plurality of disease
4 specific marker sequences, special significance was given to
5 markers which were evidentiary of a particular disease state
6 or with conditions associated with Syndrome-X. Syndrome-X is
7 a multifaceted syndrome, which occurs frequently in the
8 general population. A large segment of the adult population
9 of industrialized countries develops this metabolic syndrome,
10 produced by genetic, hormonal and lifestyle factors such as
11 obesity, physical inactivity and certain nutrient excesses.
12 This disease is characterized by the clustering of insulin
13 resistance and hyperinsulinemia, and is often associated with
14 dyslipidemia (atherogenic plasma lipid profile), essential
15 hypertension, abdominal (visceral) obesity, glucose
16 intolerance or noninsulin-dependent diabetes mellitus and an
17 increased risk of cardiovascular events. Abnormalities of
18 blood coagulation (higher plasminogen activator inhibitor
19 type I and fibrinogen levels), hyperuricemia and
20 microalbuminuria have also been found in metabolic syndrome-
21 X.

22 The instant inventors view the Syndrome X continuum in
23 its cardiovascular light, while acknowledging its important
24 metabolic component. The first stage of Syndrome X consists

1 of insulin resistance, abnormal blood lipids (cholesterol and
2 triglycerides), obesity, and high blood pressure
3 (hypertension). Any one of these four first stage conditions
4 signals the start of Syndrome X.

5 Each first stage Syndrome X condition risks leading to
6 another. For example, increased insulin production is
7 associated with high blood fat levels, high blood pressure,
8 and obesity. Furthermore, the effects of the first stage
9 conditions are additive; an increase in the number of
10 conditions causes an increase in the risk of developing more
11 serious diseases on the Syndrome X continuum.

12 A patient who begins the Syndrome X continuum risks
13 spiraling into a maze of increasingly deadly diseases. The
14 next stages of the Syndrome X continuum lead to overt
15 diabetes, kidney failure, and heart failure, with the
16 possibility of stroke and heart attack at any time. Syndrome
17 X is a dangerous continuum, and preventative medicine is the
18 best defense. Diseases are currently most easily diagnosed in
19 their later stages, but controlling them at a late stage is
20 extremely difficult. Disease prevention is much more
21 effective at an earlier stage.

22 Subsequent to the isolation of particular disease state
23 marker sequences as taught by the instant invention, the
24 promulgation of various forms of risk-assessment tests are

1 contemplated which will allow physicians to identify
2 asymptomatic patients before they suffer an irreversible
3 event such as diabetes, kidney failure, and heart failure,
4 and enable effective disease management and preventative
5 medicine. Additionally, the specific diagnostic tests which
6 evolve from this methodology provide a tool for rapidly and
7 accurately diagnosing acute Syndrome X events such as heart
8 attack and stroke, and facilitate treatment. As an additional
9 concept, the particular marker may be further validated by
10 recognition of the corresponding autoantibody.

11 In order to purify the disease specific marker and
12 further characterize the sequence thereof, additional
13 processing was performed.

14 For example, Serum (20 ml) was (diluted 5-fold with
15 phosphate buffered saline) concentrated by centrifugation
16 through a YM3 MICROCON spin filter (Amicon) for 20 min at
17 10,000 RPM at 4°C in a Beckman MICROCENTRIFuge R model bench
18 top centrifuge. The filtrate was discarded and the retained
19 solution, which contained the two peptides of interest, was
20 analyzed further by tandem mass spectrometry to deduce their
21 amino acid sequences. Tandem mass spectrometry was performed
22 at the University of Manitoba's (Winnipeg, Manitoba, Canada)
23 mass spectrometry laboratory using the procedures that are
24 well known to practitioners of the art.

1 In accordance with various stated objectives of the
2 invention, the skilled artisan, in possession of the
3 specifically isolated proteomic material, would readily carry
4 out known techniques in order to raise purified biochemical
5 materials, e.g. monoclonal and/or polyclonal antibodies,
6 which are useful in the production of methods and devices
7 useful as point-of-care rapid assay diagnostic or risk
8 assessment devices as are known in the art.

9 The specific proteomic materials which are analyzed
10 according to the method of the invention are released into
11 the circulation and may be present in the blood or in any
12 blood product, for example plasma, serum, cytolyzed blood,
13 e.g. by treatment with hypotonic buffer or detergents and
14 dilutions and preparations thereof, and other body fluids,
15 e.g. CSF, saliva, urine, lymph, and the like. The presence
16 of each proteomic material marker is determined using
17 antibodies specific for each of the markers and detecting
18 specific binding of each antibody to its respective marker.
19 Any suitable direct or indirect assay method may be used to
20 determine the level of each of the specific markers measured
21 according to the invention. The assays may be competitive
22 assays, sandwich assays, and the label may be selected from
23 the group of well-known labels such as radioimmunoassay,
24 fluorescent or chemiluminescence immunoassay, or immunoPCR

1 technology. Extensive discussion of the known immunoassay
2 techniques is not required here since these are known to
3 those of skilled in the art. See Takahashi et al. (Clin Chem
4 1999;45(8):1307) for S100B assay.

5 A monoclonal antibody specific against the proteomic
6 material sequence isolated by the present invention may be
7 produced, for example, by the polyethylene glycol (PEG)
8 mediated cell fusion method, in a manner well-known in the
9 art.

10 Traditionally, monoclonal antibodies have been made
11 according to fundamental principles laid down by Kohler and
12 Milstein. Mice are immunized with antigens, with or without,
13 adjuvants. The splenocytes are harvested from the spleen for
14 fusion with immortalized hybridoma partners. These are
15 seeded into microtitre plates where they can secrete
16 antibodies into the supernatant that is used for cell
17 culture. To select from the hybridomas that have been plated
18 for the ones that produce antibodies of interest the
19 hybridoma supernatants are usually tested for antibody
20 binding to antigens in an ELISA (enzyme linked immunosorbent
21 assay) assay. The idea is that the wells that contain the
22 hybridoma of interest will contain antibodies that will bind
23 most avidly to the test antigen, usually the immunizing
24 antigen. These wells are then subcloned in limiting dilution

1 fashion to produce monoclonal hybridomas. The selection for
2 the clones of interest is repeated using an ELISA assay to
3 test for antibody binding. Therefore, the principle that has
4 been propagated is that in the production of monoclonal
5 antibodies the hybridomas that produce the most avidly
6 binding antibodies are the ones that are selected from among
7 all the hybridomas that were initially produced. That is to
8 say, the preferred antibody is the one with highest affinity
9 for the antigen of interest.

10 There have been many modifications of this procedure
11 such as using whole cells for immunization. In this method,
12 instead of using purified antigens, entire cells are used for
13 immunization. Another modification is the use of cellular
14 ELISA for screening. In this method instead of using
15 purified antigens as the target in the ELISA, fixed cells are
16 used. In addition to ELISA tests, complement mediated
17 cytotoxicity assays have also been used in the screening
18 process. However, antibody-binding assays were used in
19 conjunction with cytotoxicity tests. Thus, despite many
20 modifications, the process of producing monoclonal antibodies
21 relies on antibody binding to the test antigen as an
22 endpoint.

23 The purified monoclonal antibody is utilized for
24 immunochemical studies.

1 Polyclonal antibody production and purification
2 utilizing one or more animal hosts in a manner well-known in
3 the art can be performed by a skilled artisan.

4 Another objective of the present invention is to provide
5 reagents for use in diagnostic assays for the detection of
6 the particularly isolated proteomic materials of the present
7 invention.

8 In one mode of this embodiment, the proteomic materials,
9 e.g. the disease specific marker sequences of the present
10 invention may be used as antigens in immunoassays for the
11 detection of those individuals suffering from the disease
12 known to be evidenced by said marker sequence. Such assays
13 may include but are not limited to: radioimmunoassay, enzyme-
14 linked immunosorbent assay (ELISA), "sandwich" assays,
15 precipitin reactions, gel diffusion immunodiffusion assay,
16 agglutination assay, fluorescent immunoassays, protein A or G
17 immunoassays and immunoelectrophoresis assays.

18 According to the present invention, monoclonal or
19 polyclonal antibodies produced against the isolated proteomic
20 materials of the instant invention are useful in an
21 immunoassay on samples of blood or blood products such as
22 serum, plasma or the like, spinal fluid or other body fluid,
23 e.g. saliva, urine, lymph, and the like, to diagnose patients
24 with the characteristic disease state linked to said marker

1 sequence. The antibodies can be used in any type of
2 immunoassay. This includes both the two-site sandwich assay
3 and the single site immunoassay of the non-competitive type,
4 as well as in traditional competitive binding assays.

5 Particularly preferred, for ease and simplicity of
6 detection, and its quantitative nature, is the sandwich or
7 double antibody assay of which a number of variations exist,
8 all of which are contemplated by the present invention. For
9 example, in a typical sandwich assay, unlabeled antibody is
10 immobilized on a solid phase, e.g. microtiter plate, and the
11 sample to be tested is added. After a certain period of
12 incubation to allow formation of an antibody-antigen complex,
13 a second antibody, labeled with a reporter molecule capable
14 of inducing a detectable signal, is added and incubation is
15 continued to allow sufficient time for binding with the
16 antigen at a different site, resulting with a formation of a
17 complex of antibody-antigen-labeled antibody. The presence
18 of the antigen is determined by observation of a signal which
19 may be quantitated by comparison with control samples
20 containing known amounts of antigen.

21 All patents and publications mentioned in this
22 specification are indicative of the levels of those skilled
23 in the art to which the invention pertains. All patents and
24 publications are herein incorporated by reference to the same

1 extent as if each individual publication was specifically and
2 individually indicated to be incorporated by reference.

3 It is to be understood that while a certain form of the
4 invention is illustrated, it is not to be limited to the
5 specific form or arrangement herein described and shown. It
6 will be apparent to those skilled in the art that various
7 changes may be made without departing from the scope of the
8 invention and the invention is not to be considered limited
9 to what is shown and described in the specification and
10 drawings/figures.

11 One skilled in the art will readily appreciate that the
12 present invention is well adapted to carry out the objectives
13 and obtain the ends and advantages mentioned, as well as
14 those inherent therein. The oligonucleotides, peptides,
15 polypeptides, biologically related compounds, methods,
16 procedures and techniques described herein are presently
17 representative of the preferred embodiments, are intended to
18 be exemplary and are not intended as limitations on the
19 scope. Changes therein and other uses will occur to those
20 skilled in the art which are encompassed within the spirit of
21 the invention and are defined by the scope of the appended
22 claims. Although the invention has been described in
23 connection with specific preferred embodiments, it should be
24 understood that the invention as claimed should not be unduly

limited to such specific embodiments. Indeed, various
modifications of the described modes for carrying out the
invention which are obvious to those skilled in the art are
intended to be within the scope of the following claims.